Coherence under Constraint, a structural model of mental health

Benjamin James

January 24, 2026

Redefining mental health is not interpretive gloss or an attempt to rescue our vocabulary by reifying its assumptions in new terms. It is a statement of minimal commitments about the least one must assume in order to speak coherently about minds, meaning, and breakdown without covertly reintroducing metaphysics whenever explanation runs thin. The following axioms are intentionally sparse because surplus premises do not remain surplus; they calcify into foundations, and foundations tend to be defended long after their use has expired. Each axiom is therefore not a preference but a constraint on what may legitimately be posited in subsequent analysis, and they are arranged so that later claims are disciplined by earlier ones rather than floating free as optional rhetoric.

Axiom I asserts constraint primacy. *All cognition, behavior, and meaning arise under constraint*. This denies a picture in which the mind begins as an internally self-sufficient agent and then encounters external limits as contingencies. Constraint is not external interference with an otherwise autonomous psyche; it is the enabling condition of any psyche whatsoever. Material limits, perceptual bandwidth, social enforcement, temporal finitude, and energetic cost do not merely restrict what a system does; they constitute what a system can be. From this follows that any explanation which treats mental phenomena as if they can be generated independently of constraint is structurally compelled to fabricate internal causes to fill its explanatory vacuum. If one subtracts real shaping forces, intelligibility must be preserved by narration, and narration, in such cases, is hallucination at the level of theory.

Axiom II defines compression necessity. *Finite systems must compress reality in order to act*. Constraint primacy entails finitude, and finitude entails selective reduction. A system with limited bandwidth cannot carry the full complexity of its environment forward into action; it must perform a lossy transformation that preserves only what it can operationalize. Symbols, narratives, identities, diagnoses, and models therefore function not as representations in the strict sense but as compressions, procedures for reducing dimensionality while retaining action-relevant structure. The loss is not accidental but constitutive of meaning's utility. A map that does not omit variables would not be a map. Likewise, a symbol that does not discard would not be a symbol but an unmanageable

duplication. This implies failure begins not with loss but with the denial of loss, that is, with the assumption that the compression is equivalent to what it compresses.

Axiom III makes this boundary explicit with symbolic non-equivalence. *No symbolic structure is equivalent to the reality it compresses*. This axiom functions as a prohibition against representational inflation, or our tendency to treat a successful compression as though it were a mirror of the world rather than a situated tool. Symbols do not exhaust their referents, remain valid across contexts, or self-correct in the absence of feedback, because their very construction is purpose-relative and constraint-relative. Their success is local, and when success is misread as universality, the symbol begins to demand that reality conform to it. The resulting rigidity is not an epistemic virtue but a failure mode. When symbols are taken as mirrors, adaptation becomes definitionally unnecessary, because the system believes it already possesses what it is meant to track.

Axiom IV defines coherence as viability. *A system is coherent if it remains viable across changing constraints*. This is a deliberate disambiguation of coherence from traditional evaluative categories. Coherence is not truth, internal consistency, moral correctness, or psychological wholeness, because all of those can be achieved in ways that are brittle, parasitic, or self-sealing. Coherence names a functional relation. It is the continued alignment between a system's compressions and the constraints they must track well enough to permit persistence. A system may remain viable while holding false beliefs, incomplete models, or provisional identities, so long as those compressions continue to support workable action within its relevant constraint field. Therefore, coherence cannot be reduced to legibility or correctness. Evaluation must be conducted at the level of persistence under constraint rather than at the level of narrative satisfaction or theoretical elegance.

Axiom V reframes repression as structural suppression. *When a behavior, perception, or signal is persistently punished, it is suppressed, not eliminated*. This denies that suppression requires a hidden chamber, buried content, or some subterranean ontology. Punishment operates by changing the cost structure of expression; it removes certain possibilities from overt enactment because their enactment becomes predictably too costly. Suppression does not erase the underlying dynamics that generated the signal in the first place. The signal remains operative but deprived of direct channels; thus, it reappears indirectly through distortion, substitution, or leakage. The implication is methodological as much as theoretical. Symptoms should not be treated as messages with latent content requiring hermeneutic decoding, but as predictable side effects of constrained option space. They indicate where suppression load is being carried, not where hidden meaning is waiting to be recovered.

Axiom VI follows from finitude again, non-conscious governance. *Most system-regulating processes operate outside symbolic awareness, not because they are occluded behind a psychic veil, but because symbolization is costly, articulation lags action, and awareness is bandwidth-limited*. Any system that attempts to render all experience explicitly would lose the speed and distribution necessary for viability; it would convert real-time control into retrospective narration. Much of what governs behavior therefore remains non-symbolic by necessity. There is no privileged "inner truth" narrative interpretation can excavate if only one digs deeply enough. What exists is untracked structure; processes that operate because they must, regardless of whether the system can narrate them. Our romanticization of hidden depth is replaced by a disciplined accounting of what is active and what is being represented.

Axiom VII names symbolic drift. *All symbolic systems drift over time away from the constraints that originally stabilized them*. Drift is not an anomaly but a structural consequence of the differential speeds at which environments change and symbols stabilize. Environments are dynamic; symbols, once successful, tend to harden, precisely because success produces trust, trust produces defense, and defense delays revision. Drift accumulates most efficiently when a system appears to be working, because visible failure is the condition under which revision becomes socially and psychologically permissible. This means crises are not abnormal but predictable; they occur when drift reaches a threshold at which existing compressions can no longer accommodate current constraints without strain. Stability is therefore not the default state of symbolic life; it is a temporary alignment maintained only by ongoing responsiveness.

Axiom VIII treats breakdown as feedback. *Experiences labeled disorder, chaos, pathology, or hallucination are often signals of symbolic misfit, not intrinsic failure*. This does not deny that breakdown can be harmful, nor does it romanticize it as revelation. It specifies its informational status within this model. Breakdown occurs when compression exceeds tolerance, revision is blocked, and enforcement replaces adaptation; under such conditions, the system continues to operate with symbols that no longer track what they must, and the resulting misalignment manifests as destabilization. Pathologizing breakdown without examining symbolic fit intensifies the problem, because it treats a correction signal as a defect to be suppressed. Suppression of feedback preserves the misfit conditions and forces escalation with more rigidity, more suppression load, and narrower option space.

Axiom IX rejects a central governor. *There is no homunculus, ego, Self, or inner executive directing the system*. This model refuses agentic miniatures because they function as explanatory shortcuts

that replace the structural analysis of distributed processes with a single fictitious manager. What appears as agency is temporary equilibrium, a coordinated pattern sustained through ongoing effort and contingent on context. Systems appear unified when constraint conditions allow stable coordination and fragmented when those conditions shift, but neither phenomenology implies a centralized controller. The loss of control is not a fall from sovereignty, because sovereignty never existed. Our feeling of executive authority is an artifact produced by ongoing coherence, not a causal source of it.

Axiom X asserts adaptation over interpretation. Systems change through altered constraints and expanded option space, not through correct interpretation alone. Interpretation can produce narrative coherence, but narrative coherence is not the same as viability. A system may become extraordinarily articulate about its own dysfunction while remaining trapped in the constraint topology that generates it, because articulation does not itself alter costs, affordances, or available actions. Worse, insight can often deepen lock-in because it stabilizes a compelling explanatory story that substitutes for structural revision. It is not that understanding is worthless, but that its role is partial and instrumental. Understanding is valuable only insofar as it participates in reconfiguring the system's conditions of action. Otherwise, it functions as a symbolic narcotic that preserves misfit by making it narratively legible.

Axiom XI generalizes the frame through scale invariance. *These same axioms govern individuals, institutions, cultures, and intelligent systems*. Differences in substrate do not exempt systems from constraint primacy, compression necessity, drift, suppression, and misfit feedback. Mental health, bureaucracy, ideology, and AI failure differ in material implementation, but they share the same structural failure modes because they are all instances of constrained systems maintaining compressive mappings under changing conditions. Psychology is not ontologically special; it is historical and local. Its phenomena become intelligible only when situated within the broader class of coherence dynamics, rather than insulated as a private mystery of minds.

Axiom XII introduces the minimal normative criterion of ethics as coherence preservation. *An action is ethical if it increases coherence across scales without collapsing future option space*. This rejects ethical foundations rooted in rule-following, outcome maximization, or moral narratives, because each can generate local stability by externalizing costs into the future, other agents, or other scales of the system. Ethical action, within this model, is structural stewardship under uncertainty; behavior that increases viability without mortgaging adaptive capacity. The key constraint is option space, because coherence purchased by collapsing possibility is unstable by construction; it

achieves present alignment by imposing future fragility. Ethical gains must therefore preserve, rather than foreclose, the system's ability to revise itself as constraints change.

Together, these axioms form a structural mapping, not an appeal to authority and not a defense of any diagnostic canon. Diagnoses are treated neither as truths about persons nor as fictions to be dismissed, but as stabilized descriptions of recurring failure patterns in constrained systems. Our task is not to interpret them psychologically, nor to rescue them institutionally, but to translate them into the language of coherence failure. *What breaks, under what conditions, and why those breaks recur with such reliability*. This translation proceeds in two layers. First, major diagnostic clusters are mapped to specific failures of coherence as defined by the axioms already stated. Second, the institutional environments that reliably generate those failures are examined, not as accidental contributors, but as active producers of the very conditions labeled as pathology. This is not critique by rhetoric or indignation. It is an account of structural causation.

Depression, when stripped of its diagnostic reverence, is best understood as option-space collapse. Its defining feature is not sadness, low mood, or affective flattening, which are secondary phenomena, but the progressive narrowing of viable future actions. The system encounters a constraint landscape in which exploration is consistently punished, ignored, or rendered futile. Attempts at revision fail to produce meaningful change, while the cost of further attempts continues to rise. Under such conditions, the system conserves energy by reducing engagement. This conservation is often described as withdrawal, anhedonia, or lack of motivation, but structurally it is a rational response to an environment in which action no longer reliably improves outcomes. The subjective experience of hopelessness corresponds to an accurate perception that the future has become inaccessible not because the system cannot imagine alternatives, but because those alternatives are not safely enactable. Depression, thus, is not a disorder of valuation but a consequence of sustained constraint without revision pathways.

Anxiety follows a different but equally intelligible failure pattern. Here the primary issue is suppression overload under threat saturation. The system is required to monitor too many potential dangers simultaneously while being punished for visible error. Signals that might otherwise be exploratory or corrective are flagged as risky to express or act upon. The result is continuous anticipatory regulation. Internal monitoring costs escalate as the system attempts to prevent failure before it occurs, because failure carries disproportionate consequences. Racing thoughts, hypervigilance, and somatic distress are not signs of irrational fear; they are the physiological correlates of sustained over-monitoring in an environment that does not tolerate deviation. Anxiety

is the cost of successfully preventing breakdown in a system that enforces compliance more aggressively than adaptation. It persists not because the system misperceives threat, but because the threat structure is real and inescapable.

Obsessive–compulsive patterns emerge when option space narrows to a single reliably safe channel. Under conditions of high uncertainty and punitive feedback, most actions become too risky to attempt. One behavior, however arbitrary in content, remains consistently available, a ritual, check, or repetition that produces temporary relief. This behavior persists not because it is meaningful, but because it works. Repetition substitutes for adaptability when adaptability is no longer safe. Intrusive doubt arises because the system cannot trust its broader environment, and so must continually verify the one channel it can still control. Compulsion is not excess order but adaptive fixation. It is what happens when only one action still functions under constraint.

Psychosis, broadly construed, represents a collapse in constraint tracking at the level of shared symbolic reference. This is often misunderstood as a loss of contact with reality, but the more precise failure is a loss of access to collectively enforced arbitration. Symbolic systems that once mediated between private experience and public meaning no longer reliably map onto lived constraints. Correction signals from the environment are delayed, contradictory, or experienced as punitive rather than informative. Over time, the system begins to rely on private compressions to restore coherence. Delusions and hallucinations arise not as arbitrary fabrications, but as attempts to make sense of experience in the absence of trustworthy shared reference points. Psychosis is therefore not disconnection from reality per se, but disconnection from the mechanisms by which reality is socially stabilized. It often follows prolonged symbolic invalidation, where the system's reports of misfit are consistently dismissed or punished rather than integrated.

What is commonly labeled ADHD can be reframed as environmental misfit rather than intrinsic deficit. In this case, the system tracks too many signals in an environment that rewards monotonic attention and linear task structure. Distributed awareness, rapid context switching, and sensitivity to novelty are penalized, not because they are dysfunctional, but because institutional pacing and evaluation mechanisms cannot accommodate them. Option space is externally narrowed to forms of engagement that conflict with the system's natural compression style. The resulting distress, disorganization, or underperformance is then attributed to internal disorder rather than to the hostility of the environment to cognitive diversity. Pathology is a consequence of enforced misalignment.

Autism-related distress follows a similar pattern, though the failure mode centers on symbolic over-enforcement. Non-symbolic or differently-symbolized processing is punished, while verbal, social-symbolic performance is demanded as a condition of participation. Masking becomes continuous, requiring the system to suppress natural processing styles in favor of externally legible ones. Suppression load accumulates, flexibility diminishes, and exhaustion follows. Distress does not arise from autism itself, but from compulsory symbolic compliance in environments that privilege narrow forms of expression and interpretation. The system is asked to translate itself constantly into a language that does not fit its structure.

So-called personality disorders, when generalized beyond their stigmatizing labels, can be understood as cases of early compression lock-in. In environments where survival depends on rapid stabilization of a narrow strategy, certain compressions become rigid very early. These strategies remain internally coherent because they once worked under severe constraint. Revision threatens identity collapse because the compression is not merely a belief or habit, but the condition under which the system first achieved viability. Later contexts may render this strategy maladaptive, but from the system's perspective, abandoning it risks annihilation. What appears externally as stubbornness, volatility, or maladaptation is often the persistence of a once-necessary solution in the absence of safe revision pathways. These are frozen adaptations, not character defects.

This mapping makes one fact unavoidable; *mental illness is not evenly distributed because constraints are not evenly applied*. The second layer of this analysis therefore turns to the institutions that reliably produce these failure conditions. Education systems, for example, enforce standardized pacing, narrow attention norms, and early identity labeling. Divergence is punished, often before it can be articulated as difference rather than deficiency. Children learn quickly which forms of attention and expression are rewarded, and which are costly. Anxiety, learned helplessness, and the pathologization of cognitive diversity are predictable outcomes. Structurally, education optimizes for symbolic legibility and administrative manageability, not for cognitive viability.

Workplaces reproduce similar patterns at a different scale. Output is demanded without corresponding autonomy. Metrics substitute for meaning, and evaluation becomes continuous rather than situational. Precarity ensures that deviation carries existential risk, while feedback is often abstracted away from the conditions of actual work. Under such constraints, depression, burnout, compulsive overwork, and anxiety are not anomalies but expected responses. The structural error is clear; workplaces demand agency while systematically removing control.

Individuals are held responsible for outcomes they cannot meaningfully influence, a condition known to produce learned powerlessness and chronic stress.

Bureaucratic and governance systems intensify these dynamics by prioritizing symbolic compliance over material reality. Rules are enforced at a distance from their consequences, and responsibility is distributed without authority. Feedback loops are delayed or distorted, making correction difficult even when failure is evident. Punitive enforcement replaces adaptive revision, because enforcement is symbolically legible while adaptation is not. The resulting experiences of paranoia, moral injury, and disengagement are individualized as psychological problems rather than recognized as rational responses to incoherent systems. Bureaucracy rewards appearing coherent over being coherent and punishes those who attempt to expose its misfit.

Mental health institutions themselves are not exempt from this pattern. By pathologizing adaptive responses to harmful environments, they individualize systemic failure. Diagnoses become identities, and distress becomes evidence of internal defect rather than relational misfit. Treatment aims to restore functionality without altering context, stabilizing individuals just enough to return them to the conditions that produced their distress. Resistance to treatment is then interpreted as further pathology rather than as feedback about the limits of individual-focused intervention. In this way, mental health systems contribute to chronicity by enforcing narratives of disorder that foreclose structural critique.

Digital environments add another layer of constraint that is both pervasive and poorly regulated. Attention is fragmented by design, comparison is constant, and symbolic performance is incentivized. Algorithms reward extremes because extremes capture attention, and attention is the commodity being extracted. Identity instability, anxiety, depression, and dissociation are not side effects, but predictable outcomes of systems optimized for engagement. Digital platforms monetize erosion, converting instability into profit while externalizing the psychological cost onto users.

Taken together, these institutions enforce symbolic rigidity under rapidly changing constraints. They demand stability from individuals while remaining structurally resistant to revision themselves. My unifying claim is not rhetorical but empirical; *modern societies systematically generate mental illness by design*. This is not because they intend harm, but because they optimize for symbolic manageability, efficiency, and control rather than for coherence under constraint. Treatment models that ignore this fact tend to stabilize individuals just enough to re-enter harmful systems, medicalize resistance, and externalize responsibility. They address symptoms while preserving the structures that produce them.

This mapping matters because it dissolves several persistent illusions at once. It undermines social stigma by showing that distress is often a rational response to material constraint. It weakens the chemical imbalance myth by locating causation in structural conditions rather than internal deficits. It dissolves our illusion of individual defects and makes prevention visible as a design problem rather than a therapeutic one. Most importantly, it assigns responsibility where it belongs, not to isolated minds struggling to cope, but to systems that systematically violate the conditions required for mental health.

The prevailing approaches to mental health commit a category error so consistent it has become invisible. They treat failures of correspondence produced by particular constraint landscapes as if they were internal malfunctions located within individual minds. This is not merely a theoretical mistake; it is a practical one with predictable consequences. When breakdown is framed as an internal defect, intervention targets the individual's internal state, their mood, thought content, affect regulation, and verbal compliance, while leaving the structures that generated the failure intact. The result is an endless cycle of symptom management in which the underlying conditions remain untouched. This is not because clinicians or institutions are malicious or negligent, but because the diagnostic framing itself directs attention inward, away from environments, incentives, and enforcement regimes that shape what forms of health are possible in the first place.

This misclassification reliably produces interventions optimized for stabilization rather than repair. Stabilization aims to reduce visible volatility with fewer crises, fewer disruptions, fewer complaints, and fewer deviations from expected behavior. It is legible, measurable, and administratively tractable. Repair, by contrast, aims to restore adaptive coupling between a system and its constraints by reopening option space and reestablishing functional feedback. Repair is slower, riskier, and harder to quantify. It often increases short-term variability as suppressed signals re-enter the system and new strategies are tested. Modern systems overwhelmingly prefer stabilization because it aligns with institutional incentives. It reduces liability, restores predictability, and produces short-term indicators of success. This tension between stabilization and repair is therefore not accidental but structural. A system optimized for legibility will consistently choose stabilization, even when stabilization entrenches the very conditions that made repair necessary.

Therefore, recovery must be redefined. Recovery cannot mean a return to baseline, because baseline itself may have been a coerced equilibrium maintained by suppression rather than a state of genuine viability. Many individuals function for long periods by narrowing themselves to fit hostile constraints, carrying high suppression loads while appearing outwardly stable. When that configuration

collapses, “returning to baseline” means reinstating the very dynamics that produced collapse. Recovery, instead, must be understood as the restoration of viable degrees of freedom, or the return of multiple safe action pathways, multiple interpretive pathways, and the capacity to revise without catastrophic penalty. A recovered system is not one that feels good or behaves normally, but one that can move, adjust, and respond without immediately triggering punishment or collapse.

Three conditions are necessary for recovery in this sense. The first is a reduction in punitive exposure. As long as deviation reliably produces disproportionate cost, the system cannot explore alternatives without unacceptable risk. The second is the reintroduction of agency through real control surfaces. Agency here does not mean the abstract feeling of control, but actual influence over conditions that matter, such as time, pacing, task selection, relational boundaries, and material resources. Responsibility without authority does not produce health; it produces learned powerlessness. The third condition is the reintegration of feedback without catastrophic penalty. Signals of misfit must be allowed to register as information rather than being immediately reframed as failure or pathology. Without these three conditions, recovery efforts tend to produce temporary stabilization, followed by relapse once the individual is re-exposed to unchanged constraints.

This redefinition of recovery implies a corresponding shift in how intervention is understood. Intervention, under a coherence–constraint model, is not primarily a therapeutic or interpretive activity. It is a form of constraint engineering, deliberate changes to environments, incentives, social enforcement, time structures, and material conditions that alter what actions and interpretations are possible. Interpretation and insight are not excluded, but they are demoted from primary mechanisms to supporting roles. Insight matters only insofar as it enables structural modification; when insight substitutes for modification, it becomes another layer of symbolic stabilization that preserves misfit by explaining it.

Constraint engineering can be usefully classified by what it modifies. Constraint reduction lowers the cost of failure by reducing surveillance, punitive enforcement, and precarity. When errors are survivable, exploration becomes possible. Constraint redistribution moves burden away from isolated individuals and onto systems with greater capacity to absorb variance. This may involve collective responsibility, redundancy, or institutional buffering rather than individualized coping. Constraint clarification reduces contradictory demands and symbolic ambiguity. Many systems generate distress not through overt punishment but through incompatible expectations that cannot be simultaneously satisfied. Option-space expansion increases the number of permissible strategies by adding slack, redundancy, and alternative pathways rather than forcing convergence on

a single “correct” mode of functioning. Feedback repair shortens loops, makes consequences actionable, and removes punitive distortion so that signals of misfit can guide adaptation rather than triggering suppression.

These forms of intervention are often dismissed as impractical or beyond the scope of mental health work, but they are routinely implemented in other domains, from engineering and safety design to organizational management, precisely because they increase system reliability. Their relative absence in mental health reflects not their infeasibility but the persistence of a category error that locates failure inside individuals rather than in the interaction between individuals and their environments.

If these interventions are so clearly aligned, why are they resisted? The answer lies in institutional incentive structures. Institutions are not neutral containers; they are systems with their own compressions, metrics, and drift dynamics. They tend to select for symbolic legibility, risk diffusion, and liability minimization. Legibility requires standardization and simplification, which conflicts with the accommodation of cognitive diversity and contextual nuance. Risk diffusion favors procedures that distribute responsibility without distributing authority, protecting the institution at the expense of the individual. Liability minimization encourages enforcement-heavy approaches that reduce visible variance, even if they increase hidden instability.

In enforcement-heavy systems, deviation is treated as a governance problem rather than as feedback. Distress, confusion, and noncompliance are reframed as individual shortcomings that must be corrected, rather than as signals that the system itself may be misaligned with its environment and participants. This reframing allows institutions to preserve their existing structures while externalizing their cost of misfit. It also accelerates symbolic drift. As institutions become increasingly self-referential, they rely more heavily on internal metrics and narratives to evaluate success, and less on external feedback. When feedback does arrive, most often in the form of burnout, attrition, or crisis, it is treated as noise or pathology rather than as information.

This dynamic is not limited to overtly authoritarian systems. Even well-intentioned institutions drift toward self-protection because adaptation carries risk. Revising policies, redistributing authority, or tolerating error threatens established power relations and exposes the institution to uncertainty. It is therefore rational, from the institution’s perspective, to suppress feedback and stabilize appearances. Mental health becomes a downstream casualty of this logic, because the cost of maintaining institutional coherence is offloaded onto individuals who must absorb its misfit internally.

If mental health is to be taken seriously as a systemic outcome rather than an individual trait, then systems must be designed with coherence in mind. This requires articulating design principles that are structural rather than moralistic. Slack is non-negotiable. Systems without slack convert variance into pathology because they have no buffer for deviation. What could have been absorbed as noise becomes a crisis requiring suppression. Local autonomy is required. When individuals are held responsible for outcomes they cannot influence, suppression and disengagement follow. Error tolerance must be engineered rather than rhetorically endorsed. Punishment-first environments reliably generate anxiety and compulsion because they force systems into anticipatory regulation. Diversity of cognitive styles must be structurally accommodated rather than nominally celebrated. Normalization as policy manufactures disorder by defining a narrow band of acceptable functioning and penalizing everything outside it. Feedback must be fast, honest, and non-punitive. When feedback is delayed or distorted by fear of punishment, systems learn to hide problems rather than solve them. Finally, symbols must remain revisable. Policies, roles, and narratives must be treated as hypotheses subject to update, not doctrines to be defended.

These principles are not mental health recommendations in any conventional sense. They are general conditions for all adaptive systems, whether those systems are organizations, technologies, communities, or minds. Mental health emerges as one manifestation of their presence or absence, not as a special domain requiring separate rules. A system that satisfies these conditions will tend to produce fewer failures labeled as mental illness, not because it treats individuals better, but because it makes fewer impossible demands.

Recognizing this shifts how measurement itself must be approached. Current measurement regimes prioritize what is administratively convenient with symptom inventories, compliance rates, productivity metrics, attendance, and utilization statistics. These measures are attractive because they are easy to collect, compare, and report. Unfortunately, they are poorly correlated with viability. A system can score well on symptom reduction while steadily eroding option space, increasing suppression load, and lengthening feedback loops. By the time failure becomes visible, it is often too late to intervene without drastic measures.

A coherence-oriented measurement regime would instead track indicators that reflect the system's adaptive capacity. Option-space breadth asks how many viable strategies an individual can enact without incurring disproportionate penalty. Suppression load estimates how much of a system's functioning is devoted to inhibition, masking, or self-monitoring rather than to productive engagement. Feedback latency measures the time between action and actionable consequence;

long latencies encourage guesswork and anxiety. Punishment density captures how frequently and severely deviation is penalized, a strong predictor of suppression and rigidity. Revision cost assesses how expensive it is, materially, socially, and psychologically, to change roles, routines, relationships, or narratives. These metrics are not metaphysical; they can be operationalized and observed. Institutions rarely track them, and when they do, they optimize them in the wrong direction.

The ethical implications of this analysis are unavoidable. Once mental illness is understood as a predictable output of coercive, rigid, drifted symbolic systems, responsibility cannot remain individualized without becoming incoherent. To continue assigning blame or burden to individuals while preserving the structures that generate distress is not merely unjust; it is logically incoherent. Ethics must be restated as preservation across scales. Interventions that stabilize one scale by collapsing another, whether stabilizing institutions by destabilizing individuals or stabilizing the present by foreclosing future possibility, are ethically untenable. They defer crises rather than resolving them.

A society that wants to improve mental health must redesign its constraint environments. It must accept the short-term cost of increased variability in exchange for long-term viability. Otherwise, it must acknowledge that it is choosing pathology as a governance tool and using suppression, normalization, and individualization of failure to maintain symbolic order in the face of changing constraints. There is no middle ground. The question is not whether mental illness will continue to appear, but whether it will be treated as feedback to be integrated or as noise to be silenced.

References

James, B. (2025). A world without cause. PhilPapers. https://philpapers.org/rec/JAMAWW-2

James, B. (2025). Against control. PhilPapers. https://philpapers.org/rec/JAMACB

James, B. (2026). The Chaos Engine. PhilPapers. https://philpapers.org/rec/JAMTCE-4

James, B. (2025). The Coherence Engine. PhilPapers. https://philpapers.org/rec/JAMTCE

James, B. (2025). Divergent Modes of Perception; how “hallucination” hides our hubris. PhilPapers. https://philpapers.org/rec/JAMDMO

James, B. (2025). Normal in a Neurotypical World, or Mental Health in the Grip of the Symbolic Attractor. PhilPapers. https://philpapers.org/rec/JAMNIA

James, B. (2025). Reality’s Experiment: Neurotypical Lock-in, Neurodivergent Disruption. Kindle Direct. Audible. https://philpapers.org/rec/JAMREN